FATIGUE DESIGN IN ENGINEERING MATERIALS

PROFESSOR J. F. KNOTT
The University of Birmingham
School of Metallurgy and Materials
Edgbaston, Birmingham, B15 2TT
England

ABSTRACT

The paper treats the application of fracture mechanics to the assessment of the fatigue lives of engineering components. Applications are split broadly into *Structures* and *Machines* and attention is paid to the nature and size of any defect likely to be present. An account is given of the way in which defects can be controlled and characterised. The two applications areas are related to a Kitagawa plot, modified by the incorporation of a closure-free threshold. Here, a "fatigue limit" approach applies to design for defects smaller than 0.05mm: a "fracture mechanics" approach for defects larger than 0.5mm. Examples are given for structures and machines. Defects of intermediate size are found in powder-formed alloys and in castings, and some detailed consideration is given to these.

KEYWORDS

Fatigue, fracture mechanics, thresholds, overloads, non-destructive inspection, fatigue design curves, steels, superalloys, castings, aluminium alloys.

INTRODUCTION

The application of the principles of fracture mechanics to fatigue has attained a large measure of sophistication and maturity. Important stages can be identified, starting with the general acceptance of the Paris law, to relate the crack-growth increment per cycle, da/dN, to the stress-intensity-factor range, $\Delta K = K_{max} - K_{min}$, where K_{max} and K_{min} are, respectively, the maximum and minimum values of stress-intensity factor during the fatigue cycle. The Paris relationship (Paris and Erdogan 1963) may be written as:

$$da/dN = A\,\Delta K^m \qquad 1)$$

where A and m are constants. If equation 1) were to hold, it would follow that a graph of log (da/dN) vs. log (ΔK) would be a straight line of slope m. In practice, it is observed that such a graph contains three parts: a central ("Paris law") region, a region of high values of (da/dN), as K_{max} approaches the material's fracture toughness, K_{Ic}, and a region at low ΔK, where (da/dN) becomes vanishingly small at a "threshold" value, ΔK_{th}.

Behaviour at high K_{max} is associated with the occurrence of monotonic modes of fracture (Beevers et al 1975). In particular, it has been shown that, for a system involving a combination of striation growth and transgranular cleavage, the area fraction of cleavage, A_c, increases in a steep, more or less linear, fashion with K_{max}. By inverting the graph, it can be shown that the term $(1-A_c)$ increases in a linear fashion with the parameter $(K_{Ic}-K_{max})$. The cleavage area fraction is produced virtually instantaneously (a quarter-cycle) so that the area fraction that remains to be separated by fatigue is $(1-A_c)$. The macroscopic value of (da/dN) is apparently increased by a a factor $(1-A_c)^{-1}$, which is proportional to $(K_{Ic}-K_{max})^{-1}$ (Knott 1986). The occurrence of rather brittle monotonic modes also helps to explain some very high values of the exponent m (in equation 1) observed in early experiments. Critical work using a sensitive "potential-drop" technique to measure crack growth and employing controlled variations in stress-ratio, $R=K_{min}/K_{max}$, demonstrated that apparent increases in 'm' were, in fact, the result of small "bursts" of monotonic fracture (Ritchie and Knott 1973). Based on this concept, it was possible to produce a compilation of experimental results which showed that high m values were always associated with materials of low fracture toughness. The values of 'm' for tough materials fell within a narrow band: 2-3 (values rose to 10 for brittle material).

The form of equation 1) combined with experimental values of m has naturally led to a number of attempts to develop micro-mechanical models for fatigue-crack-growth. Some of these relate to strain-accumulation or to the accumulation of strain-energy density, but it has been shown that if a specimen containing a growing fatigue crack is given an intermediate anneal, the growth-rate immediately following the anneal is *increased*, whereas strain-accumulation would predict a decrease, or even an incubation period, whilst strain was being re-accumulated (Knott 1986). Based on the more recent experimental data, the most plausible models relate (da/dN) to the range of crack-tip opening displacement (CTOD), $\Delta\delta$. In vacuum, there is a large amount of reversibility of slip and potential re-welding of exposed surfaces, so that values of (da/dN) are low. In inert environments, gas adsorption tends to prevent rewelding, giving rise to higher growth-rates. Even higher rates can be observed if the environment is interactive (in terms of corrosion or oxidation), but, in all cases, the growth-rate is a function of $\Delta\delta$, because $\Delta\delta$ controls the amount of clean area per cycle, which is exposed and available for interaction.

Under monotonic loading in plane strain, the value of CTOD, δ, is given by $\delta = K^2/2\sigma_Y E$, where σ_Y is the uniaxial tensile yield stress and E is Young's Modulus. One interpretation of $\Delta\delta$ is that it is the CTOD range, given by $\Delta\delta = \delta_{max} - \delta_{min}$, where δ_{max} relates to K_{max} and δ_{min} relates to K_{min}. In terms of ΔK ($=K_{max} - K_{min}$) and stress-ratio, R, it is possible to modify equation 1) but under conditions of fatigue, it is appropriate to replace the uniaxial yield stress by the reversed cyclically-hardened flow stress, $2\sigma_f$. Incorporating also the high K_{max} observation, it is possible to write:

$$da/dN = A\,\{(1+R)/(1-R)\}\ (\Delta K^2/4\sigma_f E)(K_{Ic} - K_{I\,max})^{-1} \qquad 2)$$

where A is a suitably modified constant. The final term bears close resemblance to the (empirical) form given by Forman et al (1967), and there is experimental evidence for the {(1+R)/(1-R)} factor (Knott 1986).

THRESHOLDS, OVERLOADS, SHORT CRACKS.

A large amount of interest has been shown in behaviour at low ΔK values, in particular with the existence and validity of a fatigue-crack "threshold", ΔK_{th}, below which the fatigue-crack growth-rate is vanishingly small (analogies with a sharp, "cut-off", fatigue limit in smooth testpieces are clear). Research has established that many experimental values of ΔK_{th} are features of the idiosyncratic ways in which a "long-crack" threshold test is carried out: typically, crack growth is started at a moderate ΔK value, of, say, 10-15 $MPam^{1/2}$, and the value of ΔK is then decreased (in decremental steps of some 5%), once the crack has grown a distance equal to four times the previous reversed plastic zone size. Such a "wind-down" sequence gives rise to the possibility of the operation of a number of (plane strain) "closure" mechanisms, of which plasticity-induced closure, roughness-induced closure, and (hydr)oxide-induced closure represent three major mechanisms (Suresh and Ritchie 1984, Kendall and Knott 1988). The effect is to reduce the effective stress-intensity-factor range, ΔK_{eff}, from (K_{max} - K_{min}) to (K_{max} - K_d) where K_d is the closure value.

Primarily, the reason for "premature" closure is that the plastic deformation in the "plastic wake" generates residual stresses which tend to "pin" the crack behind the tip. Roughness-induced closure and (hydr)oxide-induced closure tend to be significant at distances behind the tip of one or two grain diameters, but more general "plastic-wake closure" can operate at much larger distances. Cutting into the plastic wake, e.g. by spark machining, can give rise to large increases in crack growth-rate, because pinning forces are removed (James and Knott 1985).

Similar effects on crack growth can be induced by single or multiple overloads during crack propagation. If the value of K_{max} associated with the overload is not so high that monotonic fracture modes are induced, the general effect of an overload is to induce compressive residual stress, such that the subsequent value of da/dN is reduced, because the effective value of ΔK is reduced (Damri and Knott 1991). Overload effects relate to the treatment of crack growth under variable-amplitude loading.

The significance of the effects of the "wind-down" sequence in "long-crack" threshold tests is that the validity of ΔK_{th} as a design parameter is challenged. It is clear that ΔK_{th} is sensitive to stress-ratio, R, and that this can be rationalised in terms of the "plastic wake". For steel, the minimum value of ΔK_{th}, ΔK_o, is only some 2-3$MPam^{1/2}$. It is argued that many defects found in service have been generated by mechanisms which do not involve any plasticity that could give a "clamping" effect, so that the appropriate design parameter is the closure-free threshold (Kendall et al 1986). A substantial amount of research has been carried out also on "short cracks", but the amount of associated plasticity and closure effects in these experiments is not always clear. The distinction between short cracks, long cracks and associated design criteria is best discussed in terms of practical applications.

FATIGUE-IN-SERVICE: STRUCTURES AND MACHINES

It is of value to regard the majority of engineering components as parts of *structures* or *machines*. The characteristics of a structure are that it is basically static in function (e.g. a bridge or a drilling-platform), although, in fulfilling its function, it may be subjected to a variety of "live" loads (winds, waves, traffic loading); it is usually large, fabricated by welding, bolting, riveting or adhesive bonding; and is generally subjected to modest stresses: a design stress of "two-thirds yield" in a weldable, structural steel of typically 450MPa yield stress implies a design stress of 300MPa. It should be noted, however, that, even if the dynamic stress range is only 30MPa, a "closure free" threshold, $\Delta K_o = 3MPam^{1/2}$ implies that defects must be smaller than 3mm in size if any threshold concept is to feature in engineering design against fatigue in structures. The characteristics of a machine are that it is basically *dynamic* in function (e.g. an automobile engine, a pump, or an aero-engine); it is composed of an assemblage of quite highly-stressed moving parts and may range in size from the car engine to the 10m long x 2m diameter forgings used in turbo-alternators. In steam-turbines, maximum stresses are of order 500MPa; in gas turbines, they may reach 1GPa; and in undercarriage systems, they may be upward of 1.5GPa. Use is made primarily of quenched-and-tempered (QT) forging steels, a small amount of maraging steel (for the very high strength applications) and some titanium alloys (for high specific strength). The high design stresses, which are often based on fatigue data (because the function is dynamic), can be sustained only if initial defects are very small (as is not unreasonable for QT forging steels) or if surfaces are put into compression by thermo-chemical or mechanical means.

The nature and size-scale of defects are quite different for structures and for machines and it is necessary to consider these features in relation to appropriate fatigue design criteria. The approaches are conveniently summarised in terms of the so-called "Kitagawa plot"

(Kitagawa et al 1979). Here, the stress-range to be used as a design criterion is plotted, on a logarithmic scale, vs. the log of defect size. For a constant threshold value of ΔK, we have:

$$\Delta K_{th} = Y\,\sigma(a)^{1/2} \tag{3}$$

where Y is a constant. Taking logarithms, we may write

$$\log \Delta\sigma = -0.5\log a + const. \tag{4}$$

so that the line has a slope of -0.5. Below a critical value of a, the design value is taken to be the fatigue limit or endurance limit measured in a smooth-specimen test, since this value is smaller than the value of Δσ given by extrapolation of the threshold-based line. The Kitagawa plot is shown in fig. 1. Initially the results (obtained at R = 0.04) were related to a fatigue limit of just over 500MPa and to a threshold, ΔK_{th} of 4.8MPam$^{1/2}$, but it can be seen that adoption of a "closure free" value for ΔK_{th} gives a sharp, unambiguous cut-off and provides a lower-bound to the results. It also provides a lower-bound to many other results reported in the literature (Kendall et al 1986).

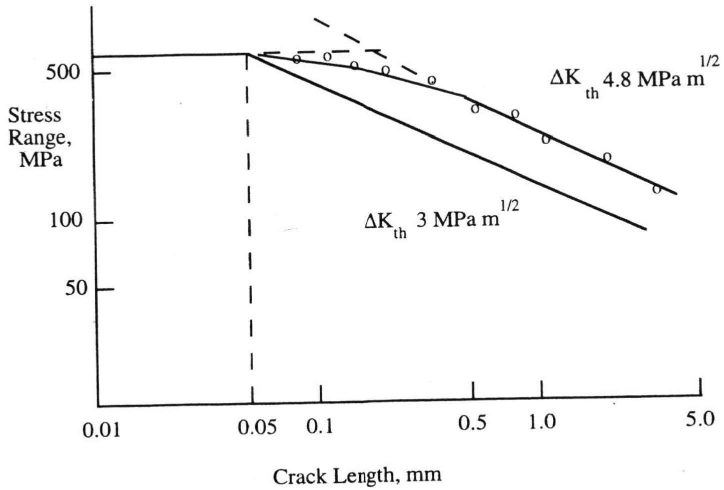

Figure. 1 Kitagawa Plot, with addition of "closure free" threshold line (after Kendall et al 1986)

The second important feature is the value of initial defect size, a, at which the design stress changes from "threshold control" to "fatigue limit control" because this interacts critically with the nature of defects in structures and in machines, with their sizes and with how they can be controlled, detected and sized. For Kitagawa's results, and his threshold value of 4.8MPam$^{1/2}$ the defect size at the transition is approx. 0.15mm and this figure, or one in the range 0.05-0.5mm, is of particular interest. Initial defects in structures are likely to result from fabrication procedures such as welding or riveting, and those of significance are likely to be upwards of 0.5mm in size. Those in wrought alloy components used in machines are usually very much smaller in size, less than 0.05mm, and may be identified with "inherent material defects", such as non-metallic inclusions, or surface defects, such as laps. The

nature and size of defects in castings and the effect of these on fatigue will be discussed in detail later. For wrought and fabricated components, however, a general guide might be that defects in *machines* (<0.05mm) are best treated using fatigue limit design, whereas those in *structures* (>0.5mm) can be treated using the closure-free threshold.

DEFECT MEASUREMENT AND CONTROL

There are basically three ways of controlling the initial defect size that enters service. For large structures, the main reliance is placed on non-destructive inspection (NDI), and, for buried defects, the only reliable technique is that of ultrasonics. For surface-breaking defects, which are often of significance in fatigue, magnetic particle inspection (MPI) or fluorescent dye techniques can give useful information, but quantification and assessment of the limitations of such techniques require further work, for example, in the examination of quenched-and-tempered steel bolts, which may contain small cracks at thread-roots, partially obscured by the oxidation associated with heat-treatment.

The efficacy of ultrasonics (U/S) NDI for the detection, location and sizing of defects in large structures has been examined in detail. In Europe, there have been a number of "Round Robin" exercises, under the auspices of PISC (Plate Inspection Steering Committee), in which various laboratories and test-houses have used their in-house procedures to detect, and locate size defects in thick-section plates containing "known" defects (this "knowledge" being confirmed by post-examination sectioning of the plates once all inspections have been carried out and logged). An even more sophisticated set of validations has been made by the U.K Inspection Validation Centre (run by AEA Technology) set up as part of the quality assurance (QA) for the Sizewell PWR pressure vessel. Here, NDI procedures, operators and equipment have been subjected to exhaustive validation exercises on representative pressure-vessel features such as "barrel" forgings, circumferential welds, nozzle/nozzle-course welded penetrations, transition course welds, upper and lower dome forgings, etc. Such "validations" give information on the accuracies to be expected from U/S NDI.

There are Physics-based limitations to what can be expected from the practical use of U/S NDI. These can be put in terms of attenuation and resolution, compared with the ease of covering large surface areas in an acceptable period of time. The usual compromise is a 25mm diameter (piezo-crystal) probe, operating at a frequency of 2 MHz. In steel, this corresponds to a wavelength of 3mm. If the frequency is increased substantially, attenuation of U/S intensity (by diffuse scattering) is such that the signal is difficult to distinguish from background noise. If the frequency is decreased, the wavelength becomes a significant fraction of the "aperture' and the "near-field" (Fresnel), diffraction-modified, intensity output is spatially distributed in a much more complicated manner than is the "far-field" (Fraunhofer) distribution, which consists of a central "lobe" of relatively focused U/S intensity, combined with well-defined "side-lobes". This is completely analogous to the optical intensity distribution emerging from a small circular hole.

Given a wavelength of 3mm, (corresponding to 2 MHz), the spatial resolution/sizing of a detected defect requires careful consideration, because the "reflector" behaves like a new "U/S aperture". Defects as small as 3mm in diameter will produce spatially "confused" diffraction peaks of a range of intensities, so that the sizing of a defect, which is usually set in terms of the "dB drop-off" (usually 20dB, although perhaps 6dB in very clean material) from the maximum signal intensity, becomes extremely susceptible to error. With a 3mm wavelength, the error in sizing is likely to be at least 3mm, more probably 5mm. These sorts of errors are observed in practice in the PISC tests and the IVC validations.

Increasing the frequency reduces the wavelength and improves resolution, but at the expense of attenuation and signal/noise ratio in thick section. Attenuation is not well formulated, but is usually put in terms of scattering, proportional to the fourth power of frequency, to the cube of grain diameter (spacing of scattering centre) and to the degree of texture. In fact, the acoustic mismatch between two grains of identical elastic properties referred to crystallographic axes, but possessing crystallographic misorientation, is likely to be far less than that at the interface between a metal matrix and a non-metallic inclusion, such as a silicate or sulphide. The cleanliness of the material is therefore paramount and there was some concern in the IVC validations that the "cleaner" Sizewell pressure vessel steel could respond to higher frequency U/S waves with less attenuation than that exhibited by the less-clean validation configurations. The gas-turbine discs in modern aero-engines are fabricated (often by powder processing) from extremely clean nickel-base superalloys, and end up as bored discs of thickness approx. 150-200mm, inner diameter 200mm and outer diameter 0.6-1.0m. Rolls-Royce employ 50MHz, 10mm diameter probes to detect defects in such discs. At this frequency, in a nickel alloy, the wavelength is reduced to approx. 0.12-0.15mm and resolution is improved accordingly, whilst the signal-to-noise ratio is acceptable, because the material is so clean in terms of inclusion content. This is despite the fact that the elastic constants for pure nickel are quite anisotropic (E in the [100] direction is 130GPa: in the [111] direction it is 287GPa), so that grain-to-grain misorientations might be thought to be significant. Certainly, there are problems with respect to U/S NDI in austenitic stainless steel weld metals (fcc like nickel) which solidify with a "cube", <100> texture.

A second possible method for defect-size control in structures is the "proof-test" or "overspeed" pre-loading. Here, the stress applied is usually some 20% higher than that which will subsequently be applied in service. If the piece does not fracture during the "proof-test" or "overspeed", there is a guarantee that no defect greater than a certain critical size exists in the piece. At normal service stress levels, this guaranteed maximum size can be used to calculate a margin or "window" for sub-critical crack growth by mechanisms such as fatigue. For conventional structural materials, these margins are not in general of such magnitude that they can provide input data to fracture mechanics predictions of life that will give reassurance. There is, however, a major effect of the "proof test"/ "overspeed" in terms of providing a residual stress field which retards early crack growth. The effect is similar to that of an overload. Analysis tends to concentrate on the "window" rather than on the "overload" effect.

For machine components, with defects less than 0.05mm in size, no conventional U/S technique is able to provide a guarantee of detection or sizing, particularly if the defect is a non-metallic inclusion, rather than an air filled crack. Under these circumstances, defect control is exerted by "process control". Metallurgical understanding is fed into features of defect production (inclusions, surface defects) such that a (low) level of (small) defects sufficient to give fatigue performance fit for service application can be identified and then reproduced, by control of processing parameters, time and again, so that every piece produced by this "indented" processing route will be known to be equally "fit for purpose".

DESIGN PHILPSOPHIES - STRUCTURES

For structures fabricated from wrought components, it is generally the case that the defects introduced by fabrication are of more significance with respect to the fatigue life than are the defects present in the original material. Two design approaches are potentially available: one, based on non-exceedance of the "closure-free" threshold; the other, making use of a form of crack growth equation such as equation 1) or 2) to calculate the number of cycles required to grow a defect from an initial size a_o to a critical size a_c. If a_o is set at a limit equal to the wavelength of 2MHz U/S waves in steel, i.e. 3mm, we have already observed that non-exceedance of a threshold value of $3MPam^{1/2}$ requires that the cyclic stress amplitude should not exceed 30MPa, although a nominal static design stress in structural steel is typically 300MPa.

If the threshold is exceeded, design must be based on crack propagation, A general form of equation 1) for steel is:

$$da/dN = 10^{-11}\ \Delta K^3 \qquad 5)$$

Using the simple form $\Delta K = \Delta\sigma(\pi a)^{1/2}$, it is possible to calculate lifetimes for different values of $\Delta\sigma$, a_o and a_c as indicated in Table I. The values chosen for $\Delta\sigma$ are 60MPa (twice the "threshold" level for a 3mm defect), 150MPa and 300MPa (the full monotonic design stress). For sensitivity, a_o is taken as 1mm, 3mm and 6mm and a_c is taken as 30mm or 60mm (related to typical value of plate thickness). The values calculated in this sensitivity study are given in Table I.

Table I - Fatigue Lifetimes for Structural Steel

a_o, mm	a_c,mm	$\Delta\sigma$,MPa	N;10^4 cycles
1	30 (60)	60	429 (458)
1	30 (60)	150	27.5 (29)
1	30 (60)	300	3.4 (3.7)
3	30 (60)	60	206 (236)
3	30 (60)	150	13 (15)
3	30 (60)	300	1.7 (1.9)
6	30 (60)	60	119 (148)
6	30 (60)	150	7.6 (9.4)
6	30 (60)	300	0.95 (1.2)

The figures in Table I may be placed in the context that a 30 year life is approx. 10^4 days. It should also be noted that an assumption of self-similar, semi-circular, surface-breaking, crack growth will increase the lifetimes in Table I by a factor of approx. $(1.5)^3$ i.e. approx. x 3.4 (it will take the 1mm defect at 60MPa below "threshold"). A detailed treatment of "thumbnail" crack growth has been given by Soboyejo et al (1990) and this has been extended to hydrogen-induced cracks of size a_o 0.5mm by Cowling and Knott (1989).

Two issues arise. The first is the occurrence of variable amplitude loading. Traditionally, this has been dealt with by means of Miner's Law summations, but it is not often emphasised that Miner's Law can be justified by means of crack growth arguments, if there is no interaction between growth under different amplitudes ("overload effects") and if the value of K_{min} does not fall below the value of K_d. Consider a number of loading "blocks" at different stress amplitudes. Suppose that the amplitude of the first block is $\Delta\sigma_1$ and that the crack grows from a_o to a_1 during the N_1 cycles for which $\Delta\sigma_1$ is applied. Let the integral of crack growth increment $\int a^{-m/2}$ da between a_o and a_1 be designated I_{o1} and let this correspond to N_1 cycles. For failure at $\Delta\sigma_1$, let the integral be denoted I_{of} and the number of cycles be $N_{f(1)}$. Then $N_1/N_{f(1)} = I_{01}/I_{of}$. Let the second block have amplitude $\Delta\sigma_2$ and let the crack grow from a_1 to a_2 in N_2 cycles. Then, integration leads to $N_2/N_{f(2)} = I_{12}/I_{of}$. For 'n' blocks, we have

$$\sum_{r=o}^{r=n} \frac{Nr}{N_{f(r)}} = \frac{1}{I_{of}} \sum_{r=o}^{r=n-1} I_{r,r+1}$$

$$= \frac{1}{I_{of}} (I_{01} + I_{12} + I_{23} ... I_{n-1,n}) \qquad 6)$$

Now, the last integral represents the last stage of growth to a_f and summation of all the integrals must be equal to I_{of}, since they simply represent incremental stages of growth from a_o to a_f. Hence the RHS (right-hand-side) of equation 6) becomes $I_{of}/I_{of} = 1$, so that the equation may be written:

$$\sum_{r=o}^{r=n} \frac{N_r}{N_{f(r)}} = 1 \qquad 7)$$

which is Miner's Law. In practice, it is known that overloads can both retard subsequent lower-amplitude propagation and give rise to increased rates due to the occurrence of monotonic fracture modes (Damri and Knott 1991),but Miner's Law serves as a good starting point. The variable amplitude loading is split up into a number of blocks and these are then subject to a Miner's Law summation.

An alternative is to subject the variable amplitude waveform to Fourier analysis and to produce a transform from the (amplitude-) time domain to the (amplitude-) frequency domain to identify "equivalence" in loading frequencies. The problem here is that if the input is not a pure sine-wave (which transforms to a single frequency), but a square-wave (or more irregular wave form), the transform contains a set of frequencies, which may, or may not, be significant with respect to fatigue crack growth. The square wave, like the amplitude distribution of light emission from an optical slit (or U/S emission from a thin, broad transducer) gives rise to a central peak (in amplitude vs. frequency or amplitude vs. distance/spread) together with a set of subsidiary peaks. In the frequency domain, this implies that the square wave is composed of a frequency spectrum and yet this concept may be irrelevant to material behaviour under a square wave "dwell " at peak load, which, in an open crack, gives the opportunity for corrosion or oxidation.

The second important point is to recognise that defects produced in fabrication are often associated with local, residual stress distributions. Defects in welded joints, such as solidification cracking or hydrogen-generated "toe" cracks, are associated with tensile residual stress fields, so that it is appropriate to assume that the full applied stress amplitude is effective. Riveted joints, on the other hand, may be associated with local compressive stresses around the rivet-hole, so that the "effective" local value of ΔK_{eff} is significantly less than the applied value, ΔK. This tends to increase the fatigue life. The pragmatic approach adopted by TWI and others has been to test simulations ("type tests") of various welded joint configurations, such as a simple butt weld, non-load-bearing fillets, load-bearing fillets, etc. The result is a set of "S-N' design curves for different "classes" of welded joints. Such design curves are shown in fig.2. It may be noted that the lifetimes calculated in Table I (give or take a factor of 3 to allow for crack shape) correspond quite well with these design curves for initial defects of size 1-3mm, giving confidence that the curves do relate back to initial defect content.

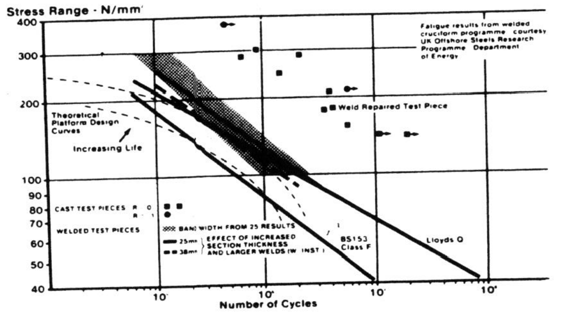

Figure. 2 Fatigue Design Curves for Welded Features (including data for a cast "node")

DESIGN PHILOSOPHIES - MACHINES

In the fatigue design of machine components, attention is drawn to the higher stress levels and smaller initial defect sizes (<0.05mm). The fatigue life requirement is often demanding: an automobile engine component, for an engine running at 60mph/4000rpm for 120,000 miles, will experience 5 x 10^8 cycles. Automobile (coil) suspension springs may have to withstand 0.25 x 10^6 cycles to "bump". Examination of the Kitagawa plot (fig. 1) indicates that, for defects <0.05mm in size, the *Engineering Design Parameter* that should be used is the fatigue limit or (5 x 10^8 cycles) endurance limit. Scientific examination of the growth of short cracks will indicate to the metallurgist how best to develop the material to improve its fatigue resistance, but the "short crack growth laws" that have been developed are, at present, not used in design.

The Kitagawa plot clearly indicates that (closure-free) threshold-based design is non-conservative for defects less than 0.05mm in size, but it is of interest to explore the consequence of using equation 5) to predict fatigue lives as a function of stress amplitude. The assumption is made that a_o = 0.03mm and that a_f = 30mm (since machine components are often of modest diameter). For $\Delta\sigma$ = 1000MPa (a typically onerous full design stress), the life is 6350 cycles; for $\Delta\sigma$ = 100MPa, it is 6.35 x 10^6 cycles and for $\Delta\sigma$ = 10MPa, it is 6.35 x 10^9 cycles. Clearly, in a machine such as an automotive engine, these calculations show that, if a_o = 0.03mm, the maximum level of (out-of-balance/vibration) fatigue stress that can be tolerated is only some 10-20MPa. It is also clear that, if the material does contain a defect of size 0.03mm, the "fatigue limit" design parameter of approx. 500MPa in the Kitagawa plot implies a life of only 5 x 10^4 cycles, or, even for a semi-circular, surface defect, somewhat less than 2 x 10^5 cycles, versus a service requirement of approx. 5 x 10^8 cycles. There are two reasons to explain the discrepancy: one , that the defects are very much smaller in size (but, note that a_o = 10μm at an amplitude of 500MPa implies only 9 x 10^4 cycles or 3 x 10^5 cycles for a semi-circular defect); two, that the effective stress intensity amplitude is significantly lower than the applied amplitude. This is effected by application of one of several surface-hardening techniques which induce compressive residual stresses in surface layers.

In steels, use is made extensively of thermo-chemical or thermal hardening techniques, such as induction-hardening, in which high-frequency currents are induced in the workpiece and are confined to surface layers by the "skin effect". This enables the surface to be austenitised, and subsequently transformed to martensite on rapid cooling, whilst the "core" remains in its state of prior heat-treatment. Case-carburising involves the diffusion of carbon into surface layers at temperatures of approx. 880°C until the outer composition is approximately eutectoid (0.8%C). This is followed by quenching. Some distortion can occur if the piece is not symmetrical, because different carbon content layers transform at different temperatures. Precision case-hardening is carried out, in Al- or Cr/V- containing steels, at temperatures of approx. 480-500°C in a cracked ammonia atmosphere, perhaps assisted by a plasma. The advantage is that the piece can be heat-treated and machined to close tolerances before nitriding. There are also purely mechanical techniques which can be applied to a wide range of materials to generate surface compression. These include shot-peening, (steels, nickel alloys, aluminium alloys; even titanium alloys, using glass beads), hammer-peening (steels, aluminium alloys), autofrettage (of steel gun-barrels), or the "ballising" of lugs (aluminium alloys). In a sense, autofrettage is simply an extreme form of an overspeed test.

A variety of techniques has been applied to establish the forms of the residual stress distributions generated by these methods. For theoretical analysis of the thermo-chemical methods, information is needed on phase-transformations and volume changes. Finite element analysis of the mechanical methods requires a knowledge of constitutive equations and of the magnitude of any Bauschinger effect, giving rise to the premature onset of reversed plasticity on unloading. Experimental techniques include strain-gauging and trepanning, but much information is derived from (high-angle) surface X-ray diffraction, combined with chemical etching. Here, it is important to recognise that the etching progressively removes the source of the residual stress, so that corrections have continually to be made, if sensible stress-distributions are to be obtained. Valuable information is also obtained from neutron diffraction. Detailed studies on the effects of "inherent" defects and residual stresses (due to quenching and tempering followed by shot-peening) on the fatigue life of automotive coil springs are being made by Todinov at Birmingham.

The general form of the stress distributions is a compression in surface layers, changing to tension in the core. A most important feature is the magnitude of the original peak tensile stress just ahead of the compression/tension interface. In some earlier work, this has been "smeared out" into a modest level of uniform tensile stress, but there is evidence that the local peak can be quite high. This may have consequences with respect to the cracking of brittle particles or the opening of voids. The growth into the piece of a fatigue-crack initiated from a surface defect is most usually analysed by loading the faces of the crack with the distribution of compressive residual stress pre-existing along the crack length and then to employ weight functions (or Green's functions) to calculate the effect of these stresses on the "negative", K_{res}, ahead of the crack tip. Sufficient "positive K", K_{max}, has to be applied to ensure that the range of stress-intensity factor, $\Delta K = K_{max} - |K_{res}|$, where $|\,|$ is "mod", or magnitude of, K_{res}, exceeds ΔK_o, the closure-free threshold. Once this is satisfied, the crack will continue to grow, in accordance with an equation of the form of 1). or 2). until the compression/tension interface is reached. A marked acceleration will then occur, and, in some cases, the value of crack length, a, at this point, virtually determines the fatigue life. The effect of surface hardening in a material containing many brittle particles (such as a particulate Al alloy/SiC MMC) is of particular interest, because there is a very real possibility that sub-surface initiation and growth could occur.

CASTINGS AND POWDER-FORMED ALLOYS.

In the above discussion, attention has been focused on the two application areas of *structures* and *machines*, loosely associated (via the Kitagawa plot, fig. 1) with values of a_o greater than 0.5mm and less than 0.05mm respectively. Defects of intermediate size can be found in different classes of engineering alloys: non-metallic inclusions in wrought alloys; porosity, oxide or intermetallics in cast products; aluminosilicate inclusions entrapped in powder-formed material.

It is of value first to address issues involved with castings. In many engineering design codes, castings are deemed to possess a degree of "unreliability", so that larger safety factors are applied, particularly when fatigue loading is involved. With respect to *structural* components, however, fig. 2 demonstrates that the fatigue performance of cast "nodes" is significantly better than that of welded joints. The reasons for this are directly related to the sizes of the initial defects The welded joints may well contain somewhere in the feature a defect of a few mm in size. The cast nodes, produced under well-controlled conditions, and subjected to exhaustive NDI, will, in general, contain defects of a much smaller size, so that the fatigue life is improved.

When castings are compared with high-quality wrought alloys for *machine* components, it is found that the castings do, in general, exhibit worse fatigue performance. Even in terms of the Kitagawa plot (fig. 1), it is apparent that the limiting design stress (based on the closure-free threshold) for a component containing a defect of 0.2mm in size (not untypical of many castings) is lower than that based on the fatigue limit/endurance limit, which is applicable to defects smaller than 0.05mm. High-quality (small defect size) castings *can* be made and such castings will possess fatigue properties as good as those associated with wrought material. Some of the most onerous high-temperature (1100°C) fatigue duty is experienced by the first-stage turbine blades in a gas-turbine aero-engine and this duty is met by the choice of nickel superalloy (directionally-solidified or single-crystal) castings.

The general principles for relating fatigue life to the initial defect size distribution in castings have been studied in detail by Taylor and Knott (1982). The application here was that of large (12m tip-to-tip diameter) marine propeller castings, made in nickel-aluminium-bronze

(NAB). For the purposes of engineering design, it was possible to rationalise design stresses in terms of the propagation (under sea-water conditions) of defects of initial size approx. 1mm. In this material, casting defects of size >0.4mm could be regarded as long cracks.

At present, Dr Jiang, working at Birmingham, is studying a comparable problem: that of the use of cast aluminium alloy "nodes" designed to fix together extruded sections of aluminium alloy to form a lightweight "space-frame" structure for automobile bodies. The design life and duty (in terms of forces) can be postulated, but these parameters must then be assessed in terms of the fatigue-life/stress-level/cross-section of casting/weight-penalty. Initially, it is important to understand the relationship between casting quality (particularly surface quality, since components are in bending) and fatigue life. In aluminium alloys, it is possible to identify four levels of quality (defect content): porosity, entrained oxide films, Fe/Si intermetallics, fine-scale microstructure. Statistical approaches help to differentiate these parameters.

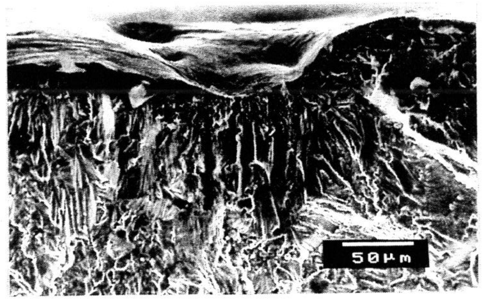

Figure 3. Fatigue-crack initiation from surface porosity in cast aluminium alloy (courtesy of Dr H Jiang)

The final case-study concerns nickel alloy gas-turbine discs, formed by a powder route, which involves atomisation, followed by hot isostatic pressing (HIPing) and forging. It is possible for particles of aluminosilicate, scoured from the refractory lining of the melting crucible or the tundish, to be entrained with the nickel alloy powder, and the size of such particles is notionally (and statistically) related to the mesh size of the sieve by which the powder particles are graded. To a large extent, it is possible to relate fatigue life to initial defect size, but, at room temperature, recognition has to be given to the effect of (thermal-contraction-induced) local residual stresses on fatigue life. At 600°C, these effects are less strong, because the residual stresses are smaller. The effect of "meso-scopic" residual stresses, induced by shot-peening, can be treated by a weight-function approach similar to that used for other residual stress distributions. H. Y. Li at Birmingham (and Xian Jiaotong) has modelled the shot-peened situation in 2-D by studying fatigue crack growth of a small defect located in a surface-breaking position in a region indented by a roller (a 2-D "ball"). The weight-function method is able to characterise fatigue-crack growth in this model system to a convincing degree of agreement.

REFERENCES

Beevers C. J, Cooke R. J, Knott J. F. and Ritchie R. O (1975). Some Considerations of the Influence of Sub-Critical Cleavage Growth during Fatigue Crack Propagation in Steels, Metal Science 9, 119-128.

Damri D. and Knott J. F. (1991). Transient Retardations in Fatigue Crack Growth Following a Single Peak Overload. Fat. Fract. Eng. Matls. Struct. 14, 709-719.

Forman R. G., Kearney V. E. and Engle R. M. (1967). Trans ASME Jnl Basic Eng, 89, 459

James M. N. and Knott J. F. (1985). An Assessment of Crack Closure and the Extent of the Short Crack Regime in Q1N (HY80) Steel. Fat. Fract. Eng. Matls. Struct, 8, 177-191.

Kendall J. M., James M. N. and Knott J. F. (1985). The Behaviour of Short Fatigue Cracks in Steels In: Proc. Europ. Conf. on Short Cracks (K. J. Miller and E. R de los Rios ed) MEP 241-248.

Kendall J. M. and Knott J. F. (1988). Near Threshold Fatigue-Crack Growth in Air and Vacuum ASTM STP 924, 103-114.

Kitagawa H., Takashima S., Suh C. M., and Miyashita. S. (1979). Quantitative Analysis of Fatigue Process - Microcracks and Slip Lines under Cyclic Strains ASTM STP675, 420-449.

Knott J.F. (1986). Models of Fatigue Crack Growth In: *Thirty Years of Fatigue Crack Growth* (R. H. Smith ed) Pergamon 31-52.

Paris. P. and Erdogan F. (1963) Trans ASME Jnl. Basic Eng. 85.

Ritchie R. O. and Knott J. F. (1973). Mechanisms of Fatigue Crack Growth in Low Alloy Steel. Acta. Met. 21, 639-652.

Suresh S. amd Ritchie R. O (1984) In: Fatigue Crack Growth Threshold Concepts (D. L Davidson and S. Suresh ed). TMS-AIME, 227-239.

Taylor D. and Knott J. F. (1982). Growth of Fatigue Cracks from Casting Defects in Nickel-Aluminium Bronze, Metals Technology 9, 221-232.